

Universidade de Brasília
Faculdade de Ciência da Informação (FCI)
Curso de Graduação em Biblioteconomia

BIBLIOTECA DIGITAL: UM MODELO APLICADO À CENTROS DE PESQUISA

Jaqueline Taketsugu Alves da Silva (09/0026403)

Brasília
1º/2013

Jaqueline Taketsugu Alves da Silva

BIBLIOTECA DIGITAL: UM MODELO APLICADO À CENTROS DE PESQUISA

Monografia apresentada ao curso de Biblioteconomia da Faculdade de Ciência da Informação da Universidade de Brasília como requisito à obtenção de grau de Bacharel em Biblioteconomia

Universidade do Brasil – UnB

Faculdade de Ciência da Informação – FCI

Orientador: Mamede Lima-Marques

Brasília

2013

Jaqueline Taketsugu Alves da Silva

Biblioteca digital para um centro de pesquisa/ Jaqueline Taketsugu Alves da Silva . – Brasília, 2013-

75 p. : il. ; 30 cm.

Orientador: Mamede Lima-Marques

Monografia (Graduação) – Universidade do Brasil – UnB

Faculdade de Ciência da Informação – FCI , 2013.

1. biblioteca sem paredes. 3. biblioteca virtual. 4. biblioteca digital. 5. biblioteca eletrônica. 6. tratamento da informação. 7. indexação. 8. catalogação. 9. classificação. I. Mamede Lima-Marques. II. Universidade de Brasília. III. Faculdade de Ciência da Informação. IV. Biblioteca digital: um modelo aplicado à centros de pesquisa

CDU 02:141:005.7



Título: Biblioteca digital: um modelo aplicado a centros de pesquisa.

Aluna: Jaqueline Taketsugu da Silva.

Monografia apresentada à Faculdade de Ciência da Informação da Universidade de Brasília, como parte dos requisitos para obtenção do grau de Bacharel em Biblioteconomia.

Brasília, 17 de setembro de 2013.

Mamede Lima-Marques - Orientador
Professor da Faculdade de Ciência da Informação (UnB)
Doutor em Ciência da Computação

Fernando César Lima Leite – Membro
Professor da Faculdade de Ciência da Informação (UnB)
Doutor em Ciência da Informação

André Henrique Siqueira – Membro externo
Doutor em Ciência da Informação

Dedico este trabalho aos meus queridos pais.

Agradecimentos

Agradeço aos meus pais pelo apoio para que toda a minha graduação fosse possível.

Ao Ravi e à Hellen por toda paciência e apoio para que este trabalho se concretizasse.

Ao professor Mamede por todo o aprendizado que me proporcionou e pela oportunidade de fazer parte da equipe do CPAI.

Às meninas Jaqueline, Isadora e Stephanie pelas discussões sobre biblioteconomia que foram importantes para o desenvolvimento da minha visão crítica da área e sem as quais o projeto da biblioteca digital não estaria onde está.

E por fim ao Lauro por ter me salvado tantas vezes com o \LaTeX .

Resumo

Em contraste com as bibliotecas tradicionais estão as bibliotecas sem paredes. Estas bibliotecas são marcadas pelo uso das tecnologias para prestar todos os seus serviços podendo ser acessadas pela internet de qualquer lugar à qualquer hora. O presente estudo tem como objetivo apresentar o processo de implantação de uma biblioteca digital em um centro de pesquisa científico. Por meio de uma revisão de literatura, se busca envolver conceitos das bibliotecas tradicionais e bibliotecas sem paredes englobando as bibliotecas virtuais, digitais e eletrônicas. Também são levantados conceitos relacionados à organização da informação. Como resultado, tem-se o relato do caso Biblioteca Digital do CPAI, explorando as experiências com as ferramentas testadas, a classificação desenvolvida e a indexação utilizada na biblioteca digital.

Palavras-chaves: biblioteca sem paredes. biblioteca tradicional. biblioteca virtual. biblioteca digital. biblioteca eletrônica. tratamento da informação. indexação. catalogação. classificação.

Abstract

In contrast with the traditional libraries there are the libraries without walls. These libraries are defined by the use of technology in order to give all their services, they can be accessed by the internet from anywhere at any time. The objective of this study is to show the process of implementation of a digital library into a scientific research center. Through a bibliographic review this study makes a relation between the concepts of the traditional libraries with the libraries without walls covering the virtual, digital and electronic libraries. Concepts regarding information organization are also discussed. As a result, there is the report of the case Digital Library of CPAI exploring the experiences with the tested tools, the developed classification and the indexation used in the digital library.

Key-words: library without walls. traditional library. virtual library. digital library. electronic library. treatment of information. indexing. Cataloguing. classification.

Lista de tabelas

Tabela 1 – Campos básicos do formato MARC 21	41
Tabela 2 – Elementos metadados Dublin Core	42
Tabela 3 – Amostra metadados BibTeX	44
Tabela 4 – Principais classes da CDD e CDU	47
Tabela 5 – Requisitos mínimos e justificativa	56
Tabela 6 – Requisitos desejáveis e justificativa	56
Tabela 7 – Metadados Dublin Core e Dspace	57
Tabela 8 – Metadados BibTeX e Dspace	59
Tabela 9 – Classes da CDU para o projeto	61

Lista de abreviaturas e siglas

AACR	Código de Catalogação Anglo-Americano
BD CPAI	Biblioteca Digital do CPAI
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CPAI	Centro de Pesquisa em Arquitetura da Informação
IFLA	Federação Internacional de Associações e Instituições Bibliotecárias
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
ISBD	Descrição Bibliográfica Internacional Normalizada
MARC	Machine-Readable Cataloging
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
RIEC	Reunião Internacional de Especialistas em Catalogação

Sumário

Introdução	19
I Preparação da pesquisa	21
1 Requisitos da pesquisa	23
1.1 Objetivos	23
1.2 Justificativa	23
1.3 Problema da pesquisa	24
1.4 Metodologia Científica	24
II Revisão de Literatura	27
2 Bibliotecas sem paredes	31
3 Tratamento da informação	37
3.1 Tratamento descritivo	37
3.1.1 Catalogação	37
3.1.1.1 Parâmetros utilizados na catalogação	38
3.1.1.2 Metadados	39
3.2 Tratamento temático	43
3.2.1 Classificação	44
3.2.1.1 CDD	46
3.2.1.2 CDU	46
3.2.2 Indexação	47
III Resultados	49
4 Introdução	51
5 Relato do caso	53
5.1 Sobre o CPAI	53
5.2 Surgimento do projeto	54
5.3 Levantamento de requisitos	55
5.4 Experiência com a primeira ferramenta	55
5.5 Experiência com a segunda ferramenta	59

5.6	Experiência com a classificação	61
5.7	Experiência com a indexação	62
Considerações finais e trabalhos futuros		65
Referências		67
Apêndices		69
APÊNDICE A Classificação BD CPAI		71

Introdução

As bibliotecas podem ser vistas como um local onde se acessa informações dos mais variados temas e nos mais variados suportes. Ao se pensar em biblioteca logo vem em mente uma construção com estantes cheias de livros e mesas para estudo. Esta primeira imagem de biblioteca está relacionada com as chamadas bibliotecas tradicionais.

As formas de prestar os serviços da biblioteca evoluíram juntamente com as Tecnologias de Informação e Comunicação – TIC, um exemplo são os catálogos em linha que aposentaram os catálogos em ficha. Essa evolução das TIC mudou tanto aspectos das bibliotecas existentes até então que com o advento da internet, novos tipos de bibliotecas começaram a surgir. Em contraste com a biblioteca tradicional, com essa característica de uma estrutura concreta, estão as bibliotecas sem paredes que prestam seus serviços de forma totalmente eletrônica.

Dentre as bibliotecas sem paredes estão as bibliotecas virtuais, digitais e eletrônicas. Em muitos casos essas bibliotecas são confundidas ou tratadas como iguais, isso porque, ainda não existe um consenso sobre o que caracteriza cada uma (CUNHA, 1999).

A organização da informação dentro de uma biblioteca envolve processos de tratamento da informação que são de fundamental importância para o funcionamento desta, dessa forma, abordar os processos de catalogação, indexação e classificação é necessário para o desenvolvimento de qualquer tipo de biblioteca sendo ela a tradicional ou as sem paredes.

O relato do caso da Biblioteca Digital do CPAI será o berço para a discussões de todos os resultados deste trabalho, explorando a experiência deste projeto desde o seu início.

Este trabalho se divide em três sessões: preparação da pesquisa, revisão de literatura e resultados. Na preparação da pesquisa estão presentes a justificativa, o problema da pesquisa, os objetivos e a metodologia científica. A revisão de literatura engloba dois capítulos: bibliotecas sem paredes e tratamento da informação. Na sessão dos resultados encontram-se os capítulos: sobre o CPAI, surgimento do projeto, experiência com a primeira ferramenta, experiência com a segunda ferramenta, experiência com a classificação e experiência com a indexação. Por fim estão as conclusões juntamente com os trabalhos futuros.

Parte I

Preparação da pesquisa

1 Requisitos da pesquisa

1.1 Objetivos

Objetivo Geral

Apresentar o processo de implantação de uma biblioteca digital em um centro de pesquisa científico.

Objetivo Específico

- a) Delimitar os fundamentos teóricos necessários à biblioteca digital e tratamento da informação;
- b) Relatar um caso, no que tange as experiências relacionadas a identificação de uma ferramenta e definição de parâmetros para uma biblioteca digital e
- c) Apresentar análise crítica de resultados.

1.2 Justificativa

Segundo [Tammaro e Salarelli \(2008\)](#), o tema biblioteca digital ganha cada vez mais espaço no contexto da Ciência da Informação e isso pode ser evidenciado pelo crescimento de publicações relacionadas a esse tema. Isso mostra o crescimento das bibliotecas digitais e no interesse por elas.

Apesar desse crescimento das bibliotecas digitais [Cunha \(2008, p.6\)](#) afirma:

Infelizmente, ainda não existem, no contexto educacional brasileiro, cursos regulares que possam preparar o profissional para gerenciar, da melhor forma possível, um projeto de biblioteca digital. Os atuais gerentes deste tipo de projeto estão aprendendo com a prática!

Esta afirmação de Cunha ocorreu à cinco anos e hoje já existem algumas instituições como o IBICT que dão consultorias para projetos de bibliotecas digitais, porém, ainda observa-se a importância dos projetos envolvidos com bibliotecas digitais compartilharem o seus conhecimentos para que esse aprendizado com a prática seja passado para outros grupos e os ajudem em seus projeto.

Neste estudo será relatado o caso do projeto da biblioteca digital do Centro de pesquisa em Arquitetura da Informação – CPAI. O CPAI é um centro multidisciplinar e interinstitucional que atua em pesquisa, desenvolvimento e inovação em tecnologia e ciência da informação, com ênfase em Arquitetura da Informação. Este centro de pesquisa possui uma produção documental resultado dos seus trabalhos relacionados à pesquisa científica, sendo assim, com o passar do tempo, o CPAI acumulou uma quantidade considerável de publicações. Dentre essas publicações estão teses, dissertações e artigos que, de alguma forma, estão envolvidos com o tema Arquitetura da Informação. Em muitos casos, as publicações de um pesquisador podem ajudar em novas publicações de outros pesquisadores e por isso a importância de se organizar e se disponibilizar esses documentos. A biblioteca digital surgiu como uma solução diante desta demanda de gerenciar toda a produção científica juntamente com outros documentos que possam ser de interesse do centro de pesquisa.

Compartilhar experiências não é apenas mencionar acertos é, principalmente, comentar o que não se concretizou como o esperado e, os aprendizados provenientes da necessidade de lidar com as mudanças de rumos no projeto. Ao compartilhar experiências, um projeto tem a chance de aprender com a trajetória do outro, havendo assim, um intercâmbio de conhecimentos que irá positivamente influenciar equipes de trabalhos.

1.3 Problema da pesquisa

Como ocorre o processo de implantação de uma biblioteca digital um centro de pesquisa científica?

1.4 Metodologia Científica

Classificação da metodologia

Esta pesquisa é de natureza aplicada “com o objetivo de examinar questões relativas a problemas práticos e suas potenciais soluções” (COZBY, 2003, p.24). O problema deste estudo é abordado de forma descritiva visto que o estudo “observa, registra, analisa e correlaciona fatos ou fenômenos sem manipulá-los” (CERVO; BERVIAN; SILVA, 2007, p.61). Como procedimentos técnicos utilizados estão o bibliográfico que “procura explicar um problema a partir de referências teóricas publicadas” (CERVO; BERVIAN; SILVA, 2007, p.60), comparativo “com a finalidade de verificar similitudes e explicar divergências”

(MARCONI; LAKATOS, 2011, p.92) e por fim o relato de caso que tem como objetivo apresentar um caso que neste trabalho é o da Biblioteca Digital do CPAI.

Percurso Metodológico

No percurso metodológico estão os passos que este estudo irá seguir para alcançar os objetivos.

- a) Coletar e analisar informações sobre bibliotecas: com base na literatura serão levantadas informações sobre bibliotecas tradicionais e bibliotecas sem paredes;
- b) Levantar conceitos de tratamento da informação: com base na literatura serão levantados conceitos dos processos relacionados à organização da informação (catalogação, indexação e classificação) e
- c) Apresentar o relato de caso CPAI:
 - Definir requisitos para biblioteca digital: em vista as necessidades do centro de pesquisa serão levantados os requisitos;
 - Delimitar ferramentas de biblioteca digital: analisar e levantar uma ferramenta para a biblioteca digital;
 - Apresentar a experiência com a classificação e
 - Apresentar a experiência com a indexação.

Parte II

Revisão de Literatura

Esta revisão de literatura é iniciada com um levantamento sobre as bibliotecas ditas como tradicionais e as bibliotecas sem paredes englobando as bibliotecas eletrônicas, digitais e virtuais. Serão apresentadas definições de cada biblioteca e as relações entre estas. Após abordar as referências relativas às bibliotecas serão expostos levantamentos sobre o tratamento da informação que envolvem os processos de catalogação, classificação e indexação. Esses três processos são a base da organização da informação da biblioteca e serão levantados nesta revisão de literatura para que dê insumos à apresentação do relato de caso no que tange esses processos.

2 Bibliotecas sem paredes

As bibliotecas passaram por transformações ao longo de sua história e as tecnologias impulsionaram muitas mudanças. [Landoni, Catenazzi e Gibb \(1993\)](#) apontam a história das bibliotecas dividida em três partes: biblioteca tradicional (referentes às bibliotecas existentes até a automação de bibliotecas), biblioteca moderna ou automatizada (marcadas pelo uso do computador para apoiar os processos de organização da informação) e, a biblioteca eletrônica (chamada de biblioteca do futuro que marca a disponibilização do texto completo *online*). Com base nesta abordagem, influenciada pelas tecnologias, serão abordadas cada uma dessas bibliotecas e juntamente com o que o autor chama de biblioteca eletrônica serão abordadas as bibliotecas virtuais e as digitais.

Dentre as bibliotecas existentes até a automação de bibliotecas encontra-se a biblioteca convencional, uma de suas características “é que tanto a coleção como os seus catálogos utilizam o papel como suporte de registro da informação” ([CUNHA, 2008](#), p.5). [Naves e Kuramoto \(2006, p.80\)](#) apontam a biblioteca tradicional como “a instituição social criada com a finalidade de adquirir, tratar, armazenar, disseminar e disponibilizar documentos, estes em sua materialidade física convencional, independentemente de sua forma ou suporte (livro, periódico, mapa, gravura, filme, CD-ROM, etc.)”. Estes autores, assim como Cunha, apontam os suportes dos documentos sendo de materialidade física porém, vão além do único suporte de papel, apontado por Cunha.

Nas últimas décadas, o computador passou a ser cada vez mais utilizado nas bibliotecas marcando o surgimento das bibliotecas automatizadas e “desde os anos 1970, muitas bibliotecas implantaram catálogos em linha” ([CUNHA, 2008](#), p.5). Esses catálogos em linha iniciaram uma nova forma de acesso aos registros das bibliotecas que comparados com os catálogos em ficha são de consulta mais rápida. [Marchiori \(2009, p.2\)](#) afirma que “uma série de catálogos de bibliotecas já estão acessíveis em toda a rede, para que qualquer um, de qualquer lugar do mundo, possa pesquisar”.

Esses catálogos disponíveis na *internet* possibilitam que uma biblioteca tenha facilmente acesso aos documentos de outra, dessa forma, os serviços de comutação e empréstimo entre bibliotecas ampliaram as possibilidades de acesso ao documento ([NAVES; KURAMOTO, 2006](#)). Apesar de ter ampliado o acesso aos documentos primários, “soluções como o Comut não mais atendem às necessidades de acesso imediato aos documentos que a informação em rede oportuniza.” ([LEVACOV, 1997, p.2](#)). Na comutação bibliográfica e no empréstimo entre bibliotecas existe um tempo a ser percorrido até que as bibliotecas se comuniquem, e uma possibilite o acesso ao documento a outra. Este processo demanda um

tempo que o usuário muitas vezes não está disposto a esperar. Naves e Kuramoto (2006, p.66) fazem uma paralelo entre o contexto tradicional e o contexto digital colocando em discurso o acesso:

Comparado com o contexto tradicional, o que o contexto digital significa é um meio de facilitar o acesso a coleções que já existiam há muito tempo, com variada dificuldade de acesso, mas cujas eventuais facilidades providenciadas não podem competir com as extraordinárias facilidades que a *internet* e a *web* podem propiciar.

O contexto digital é característico da fase das bibliotecas marcadas pelo acesso ao documento completo *online*. Essas bibliotecas, conhecidas como bibliotecas do futuro, englobam as bibliotecas digitais, virtuais e eletrônicas. Antecedente à apresentação destas, serão apresentados pontos importantes da relação delas com as bibliotecas tradicionais

Assim como no surgimento de outras concepções da era digital, que são recriações de ideias já estabelecidas, como é, por exemplo, o correio eletrônico, a biblioteca digital, num primeiro momento, espelha-se na biblioteca tradicional, para em seguida expandir esse conceito já consagrado através da apropriação e uso das tecnologias disponíveis (SAYÃO, 2009, p.13).

O autor mostra que a biblioteca digital surge a partir da biblioteca tradicional o que é importante para ficar claro que esse novo tipo de biblioteca não surgiu como algo totalmente novo e diferente da biblioteca tradicional.

No contexto das bibliotecas sem paredes “o conceito de ‘lugar’ torna-se secundário, tanto para bibliotecários quanto para usuários. O que é importante passa a ser o ‘acesso’ ” (LEVACOV, 1997, p.2), isso ocorre devido ao fenômeno de desterritorialização do documento:

Uma das principais consequências da implantação e desenvolvimento da plataforma *web* e da *internet* é o fenômeno da desterritorialização do documento, ou seja, o fato de passar o documento a ter sua materialidade desvinculada da forma física anterior, assumindo a forma digital que possibilita uma organização espacialmente integrada de textos, imagens e sons. Como produtos desse complexo cenário, encontra-se, na *web*, um ilimitado e cambiante universo constituído de novos documentos, os denominados objetos digitais (NAVES; KURAMOTO, 2006, p.78).

Para os usuários, não é relevante ter conhecimento do local onde os objetos digitais estão armazenados, seja em um servidor presente na sua cidade ou até fora do país, para ele o essencial é o acesso ao documento possibilitado pela *internet*.

Além das vantagens de acesso, as bibliotecas sem paredes ganharam espaço no universo das bibliotecas porque “mais e mais documentos estão sendo publicados nos formatos eletrônicos, o que exige redimensionamentos de espaço e mecanismos de tratamento, busca e disseminação destes materiais” (MARCHIORI, 2009, p.1).

Antes de entrar nos conceitos das variações de bibliotecas sem paredes é importante lembrar que esses conceitos se mesclam e “a biblioteca digital é também conhecida como biblioteca eletrônica (termo preferido dos britânicos), biblioteca virtual (quando utiliza os recursos da realidade virtual), biblioteca sem paredes e biblioteca conectada a uma rede” (CUNHA, 1999, p.258). Sendo assim, muitas vezes características de uma dessas bibliotecas podem estar presentes em outras e os conceitos acabam se confundindo.

Iniciando pela biblioteca eletrônica, Marchiori (2009, p.4) afirma:

A biblioteca eletrônica é o termo que se refere ao sistema no qual os processos básicos da biblioteca são de natureza eletrônica, o que implica ampla utilização de computadores e de suas facilidades na construção de índices on-line, busca de textos completos e na recuperação e armazenagem de registros. A biblioteca eletrônica se direcionará para ampliar o uso de computadores na armazenagem, recuperação e disponibilidade de informação, podendo envolver-se em projetos para a digitalização de livros. Haverá um uso extensivo de meios eletrônicos que ainda coexistirão com as publicações eletrônicas.

Tammaro e Salarelli (2008) partindo de uma análise onde a biblioteca eletrônica se caracteriza por utilizar insumos eletrônicos para a sua formação e prestação de serviços, apontam este tipo de biblioteca como algo mais abrangente que as bibliotecas digitais, podendo até ser incluída a biblioteca digital na biblioteca eletrônica. Martins (2002, p.6) afirma que biblioteca eletrônica é um “sistema no qual o processo básico da biblioteca é o eletrônico, ou seja, ampla utilização de máquinas, principalmente, microcomputadores” e fazendo uma paralelo com as bibliotecas digitais, estas “dispõem de todos os recursos de uma biblioteca eletrônica” (MARTINS, 2002, p.5).

Para Marchiori (2009) quando expressão “biblioteca virtual” é abordada com a característica principal de uso de recurso eletrônicos, esta acaba sendo utilizada como sinônimo de “biblioteca eletrônica”, isso porque os conceitos de biblioteca eletrônica possuem em comum a principal característica de utilização de recursos eletrônicos.

Sobre biblioteca virtual, para os autores Tammaro e Salarelli (2008) o adjetivo ‘virtual’ quer dizer que a biblioteca não existe, ou seja, a sua coleção de documentos está fora da biblioteca como espaço físico ou lógico. Um exemplo dado por estes autores para este tipo de biblioteca é a *World Wide Web* que possui uma coleção de documentos ligados em rede onde encontra-se o acesso aos conteúdos. Tammaro e Salarelli (2008, p.117) apontam mais um exemplo para este tipo de biblioteca, os catálogos virtuais que “são metacatálogos que, no momento em que o usuário faz sua solicitação, pesquisam simultaneamente em inúmeros catálogos e páginas Web em linha”. Marchiori (2009, p.4) aponta um conceito diferente de biblioteca virtual que está voltado para a questão da realidade virtual:

A biblioteca virtual é conceitualizada como um tipo de biblioteca que,

para existir, depende da tecnologia da realidade virtual. Neste caso, um software próprio acoplado a um computador sofisticado reproduz o ambiente de uma biblioteca em duas ou três dimensões, criando um ambiente de total imersão e interação. É então possível, ao entrar em uma biblioteca virtual, circular entre as salas, selecionar um livro nas estantes, “tocá-lo”, abri-lo e lê-lo. Obviamente, o único “lugar” onde o livro realmente existe é no computador e dentro da cabeça do leitor.

Entrando finalmente nas bibliotecas digitais [Naves e Kuramoto \(2006, p.80\)](#) afirmam:

Entende-se por biblioteca digital uma seleção de documentos, destinados a determinada comunidade, concebidos ou convertidos para o meio digital, preferencialmente em toda a sua integridade, disponibilizados na *internet*, desmaterializados de suas condições físicas tradicionais e constituídos de funções inteiramente novas, que lhes garantem hipertextualidade e caráter multimidiático.

Relacionando com os conceitos de bibliotecas no sentido mais amplo é importante apontar que

o conceito atual de biblioteca digital constitui um subconjunto de um conceito mais extenso de biblioteca, e não um substituto para ele. Todos os valores e funções da biblioteca continuam válidos, o que muda são os objetos físicos que formam a biblioteca e, naturalmente, o instrumental tecnológico para manipulá-los ([SAYÃO, 2009, p.12](#)).

O acesso ao documento completo *online* é uma das grandes características das bibliotecas sem paredes e na digital não é diferente, “ao invés de se acessarem fisicamente os documentos por meio de catálogos, sejam estes processados eletronicamente ou não, os objetos digitais na biblioteca digital se posicional em sua completude, diretamente no ciberespaço, juntamente com seus respectivos catálogos” ([NAVES; KURAMOTO, 2006, p.82](#)). [Sayão \(2009, p.14\)](#) aponta a mesma ideia de acesso imediato ao documento ao afirmar que “uma biblioteca digital possui e controla a informação. Ela oferece acesso à informação, e não apenas aponta para ela”.

Uma biblioteca digital possui um público definido, ela é formada a partir de uma demanda, isso porque “as bibliotecas digitais se dirigem a uma ou a um conjunto de comunidades de usuários” ([SAYÃO, 2009, p.15](#)). Então, não é apenas reunindo uma coleção aleatória sem propósito e organizar os documentos em uma ferramenta para biblioteca digital que o resultado será uma biblioteca digital. É necessário que a biblioteca tenha um objetivo para o seu acervo e um públicos alvo.

Após analisar vários conceitos de biblioteca digital, [Tammaro e Salarelli \(2008, p.123\)](#) apontam os elementos essenciais que a compõem:

- o usuário, entendido como o público em geral ou como usuário individualizado, do qual a biblioteca precisa conhecer as necessidades específicas

e as diversas atividades. Devem estar aptos a fazerem uso dos diversos serviços disponíveis;

- os conteúdos, isto é, os objetos digitais, organizados e estruturados nas coleções digitais segundo normas próprias e distribuídos em rede;
- os serviços de acesso, caracterizados por interfaces ou serviços mediados pelo pessoal bibliotecário.

Além desses elementos, os autores apontam um quarto elemento que são: os profissionais e as bibliotecas como instituições. Esse quarto elemento é devido a importância da biblioteca digital possuir profissionais para o seu tratamento e a visão desta como uma instituição ou, diretamente vinculada a ela.

Sayão (2009, p.15) também faz um apanhado de todas as características das bibliotecas digitais e expõe:

- as bibliotecas digitais são a contraparte digital das bibliotecas tradicionais e incluem materiais eletrônicos (digitais) bem como materiais impressos e ainda outros materiais – por exemplo, áudio, vídeo e objetos que não se enquadram na mídia impressa e nem podem ser disseminados em formato digital ainda;
- uma biblioteca digital possui e controla a informação. Ela oferece acesso à informação, e não apenas aponta para ela;
- uma biblioteca tem uma estrutura organizacional unificada com pontos consistentes para acesso aos dados;
- uma biblioteca digital não é uma entidade única, ela pode também oferecer acesso a materiais digitais e recursos de outras bibliotecas digitais;
- bibliotecas digitais apoiam o acesso rápido e eficiente a uma grande quantidade de fontes de informação distribuídas, porém vinculadas por links e que são plenamente integradas;
- bibliotecas digitais têm coleções que: a) são volumosas e persistentes ao longo do tempo; b) são bem organizadas e bem gerenciadas; c) contêm formatos variados; d) contêm objetos e não somente a sua representação; e) contêm objetos que não podem ser obtidos de outra forma;
- bibliotecas digitais incluem todos os processos e serviços oferecidos pelas bibliotecas tradicionais, embora esses processos tenham que ser revisados para acomodar diferenças entre mídias digitais e impressas;
- as bibliotecas digitais cumprem o paradigma do acesso onipresente, a qualquer hora e em qualquer lugar. Existe uma biblioteca onde houver um computador pessoal conectado a uma rede. As bibliotecas digitais estão sempre disponíveis;
- as bibliotecas digitais intensificam o conceito de compartilhamento de recursos provenientes das bibliotecas tradicionais;
- as bibliotecas digitais se dirigem a uma ou a um conjunto de comunidades de usuários.

3 Tratamento da informação

O tratamento da informação, ou tratamento documentário, é a fase do ciclo documentário onde é feita a organização dos documentos. Nesta fase se encontram os processos de indexação, catalogação e classificação (FUJITA et al., 2009). Para Naves e Kuramoto (2006, p.67) “o tratamento da informação é definido com a função de descrever os documentos, tanto do ponto de vista físico (características físicas do documento) quando do ponto de vista temático (ou de descrição do conteúdo)”.

O tratamento da informação se subdivide em: tratamento descritivo, onde se encontra o processo de catalogação; e o tratamento temático, ocupado pelos processos de classificação e indexação (FUJITA et al., 2009). Observa-se que o tratamento descritivo e tratamento temático apontado por Fujita é o que, respectivamente, Naves chama de descrever o documento do ponto de vista físico e descrever do ponto de vista temático.

3.1 Tratamento descritivo

Fujita et al. (2009, p.22) explicam que “o tratamento descritivo refere-se propriamente à catalogação, ou seja, à representação descritiva da forma física do documento (autor, título, edição, casa editora, data, número de páginas etc.)”. Ou seja, é nesta etapa onde os dados descritivos do documento são trabalhados.

Quando o usuário tem em mente um documento específico e sabe os dados dele, a busca será feita a partir das informações provenientes do tratamentos descritivo que aborda o processo de catalogação.

3.1.1 Catalogação

Um importante processo presente na biblioteca digital é a catalogação. Todo o tratamento descritivo será feito neste processo. Mey e Silveira (2009, p.7) defendem a catalogação como:

O estudo, preparação e organização de mensagens, com base em registros do conhecimento, reais ou ciberespaciais, existentes ou possíveis de inclusão em um ou vários acervos, de forma a permitir a interseção entre as mensagens contidas nestes registros do conhecimento e as mensagens internas dos usuários.

Partindo da definição da autora, é possível visualizar a catalogação como ponte entre o documento e o usuário. A partir da organização das mensagens codificadas do documento é possível a convergência entre o que o usuário procura, com relação aos dados descritivos do documento, e o que o documento tem a oferecer. Em um conceito mais simples, [Campello \(2006, p.57\)](#) define a catalogação como “a descrição das características dos documentos que formam as coleções das bibliotecas e bases de dados”.

O produto da catalogação é o catálogo. Dentro deste estarão presentes todos os registros da catalogação.

Catálogo é um meio de comunicação, que veicula mensagens sobre os registros do conhecimento, de um ou vários acervos, reais ou ciberespaciais, apresentando-as com sintaxe e semântica próprias e reunindo os registros do conhecimento por semelhanças, para os usuários desses acervos. ([MEY; SILVEIRA, 2009, p.12](#)).

Em uma visão mais simples, para [Fusco \(2011, p.43\)](#) “um catálogo pode ser exemplificado como um conjunto de metadados”.

A catalogação não é um processo simples, ao longo de sua história surgiram muitos parâmetros para torná-la cada vez mais eficiente. A seguir veremos dois desses parâmetros e na sequência, os metadados.

3.1.1.1 Parâmetros utilizados na catalogação

No universo da catalogação existem vários parâmetros voltados para tornar este processo cada vez mais padronizado. Antes de entrar nos metadados, serão abordados os seguintes:

- a) A AACR por ser um código de catalogação de grande influência na biblioteconomia brasileira e
- b) O ISBD por ter definidos diretrizes que influenciaram muitos códigos.

AACR

Para se alcançar um padrão na catalogação é necessário que o catalogador defina duas questões: os campos e a forma de descrição destes. O Código de Catalogação Anglo-Americano - AACR trabalha com esta segunda questão definindo regras para que seja possível uma padronização na descrição. Este código foi construído pela American Library Association – ALA em parceria com a Library Association (do Reino Unido) que resultou no código de catalogação considerado o mais conhecido e utilizado ([CAMPELLO, 2006](#)).

O AACR foi publicado no Brasil com a tradução em português no ano de 1969 e passou a ser adotado em quase todas as escolas brasileiras de biblioteconomia. Em 1978 foi publicada a segunda edição intitulada AACR2 que foi traduzida no Brasil entre 1983

e 1985. Após várias revisões na AACR2 surgiu a AACR2R (1988) que foi uma edição revista onde várias ementas e revisões ocorreram até 2005. Esse código continua sendo o mais utilizado no Brasil (MEY; SILVEIRA, 2009).

ISBD

A Descrição Bibliográfica Internacional Normalizada – ISBD surgiu pela primeira vez em 1969 quando Michael Gorman apresentou esta iniciativa à Reunião Internacional de Especialistas em Catalogação – RIEC. Esta norma visa padronizar as informações contidas na descrição bibliográfica (MEY; SILVEIRA, 2009).

A primeira das ISBDs foi a Descrição Bibliográfica Internacional Normalizada para publicações monográficas – ISBD (M) publicada pela Federação Internacional de Associações e Instituições Bibliotecárias – IFLA em 1971. Posteriormente, foram surgindo outras ISBDs especializadas chegando a sete normas diferentes. A IFLA com o intuito de reunir todas essas normas lançou uma edição preliminar consolidada em 2007 e em 2011 a edição consolidada foi publicada ¹.

A ISBD proporcionou um caminho para a padronização internacional visto que “todos os países se dispuseram a usá-la e esta aceitação internacional acarretou mudanças nos códigos de catalogação, que incorporaram essas normas em novas edições” (MEY; SILVEIRA, 2009, p.79). Até então existiam muitos códigos de catalogação sem haver uma busca consistente por um parâmetro internacional, a partir da ISBD foram definidas oito áreas de descrição com sistematização da ordem e de pontuação presente antes de cada campo, a partir disso foi possível definir internacionalmente áreas para descrição. A AACR2 foi estruturada a partir dessas áreas da ISBD.

3.1.1.2 Metadados

Mencionado anteriormente, na parte da AACR, para se alcançar um padrão na catalogação é necessário o catalogador definir questões: os campos e a forma de descrição desses campos. Os metadados irão definir esses campos onde serão preenchidos os dados do documento. Para Fusco (2011, p.42) “os metadados podem ser considerados um conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua organização, busca e recuperação”, já para Alves e Souza (2007, p.22) “os elementos de metadados têm o propósito primário de descrever, identificar e definir um recurso de informação com o objetivo de modelar e filtrar o acesso”. A busca que será feita pelo usuário partirá dos campos abarcados pelos metadados da catalogação, ou seja, só será possível buscar dados que foram preenchidos nos campos dos metadados a partir da catalogação, por isso, é necessário que os metadados abarquem todos os campos que possam fazer parte da busca do usuário. É importante que o usuário entenda o que cada

¹ Informações encontradas no site oficial da IFLA: <http://www.ifla.org/>

elemento dos metadados significam para que a busca seja feita de maneira correta. [Fusco \(2011, p.42\)](#) destaca que “o conceito de metadados não é algo novo, mas o uso desse termo em ambientes digitais e a variedade de padrões e as formas de utilização o são”.

Existem diversas formas de categorizar os metadados. Reunindo as categorizações mais defendidas [Sayão \(2010, p.5\)](#) apresenta:

- Metadados descritivos: é a face mais conhecida dos metadados, são eles que descrevem um recurso com o propósito de descoberta e identificação; podem incluir elementos tais como título, autor, resumo, palavras-chave e identificador persistente.
- Metadados estruturais: são informações que documentam como os recursos complexos, compostos por vários elementos, devem ser recompostos e ordenados. Por exemplo, como as páginas de um livro, digitalizadas separadamente, são vinculadas entre si e ordenadas para formar um capítulo.
- Metadados administrativos: fornecem informações que apoiam os processos de gestão do ciclo de vida dos recursos informacionais. Incluem, por exemplo, informações sobre como e quando o recurso foi criado e a razão da sua criação. Nessa categoria, estão metadados técnicos que explicitam as especificidades e dependências técnicas do recurso; inclui também os metadados voltados para apoio à gestão dos direitos relacionados ao recurso.

Dentro do universo dos metadados estão presentes os padrões de metadados que buscam padronizar a descrição de documentos. “Um padrão de metadados pode ser compreendido com um conjunto de elementos descritivos que segue um determinado modelo de dados com o objetivo de descrever recursos de um domínio específico”([FUSCO, 2011, p.44](#)).

Na biblioteconomia existem vários padrões de metadados, aqui serão abordados:

- a) MARC;
- b) Dublin Core e
- c) BibTeX.

MARC

Machine-Readable Cataloging - MARC é uma padrão de metadados utilizado para:

[...] identificar, armazenar, e comunicar informações bibliográficas em formato legível por máquina, de forma que diferentes computadores e programas possam reconhecer, processar e estabelecer pontos de acesso dos elementos que compõem a descrição bibliográfica ([ALVES; SOUZA, 2007, p.25](#)).

O primeiro projeto Machine-Readable Cataloging - MARC data de 1965-1966, [Fusco \(2011\)](#) afirma que este pode ser considerado o primeiro padrão de metadados do contexto da ciência da informação. Em 1968 surgiu o MARC II da Library of Congress -

LC que desenvolveu-se pela cooperação entre a LC e a British Library e serviu de base para inúmeros outros formatos, como, por exemplo, CALCO e UNIMARC (MEY; SILVEIRA, 2009).

O MARC tinha como objetivo “1) aceitação de todos os tipos de materiais; 2) flexibilidade para produção de diferentes aplicativos, além de catálogos; 3) utilização por diferentes sistemas automatizados” (MEY; SILVEIRA, 2009, p.77). O padrão MARC busca embarcar todos os dados possíveis que um documento possa possuir, dessa forma, é uma padrão muito extenso. Este formato faz a identificação dos metadados a partir de números, caracteres paraortográficos e letras que estão distribuídos entre “campos fixos e variáveis, subcampos e indicadores” (ALVES; SOUZA, 2007).

Fusco (2011) afirma que o padrão MARC mais utilizado para nortear os projetos de catálogos na área da Biblioteconomia é o denominada MARC 21 mantido pela LC²

Fusco (2011, p.49) resume as características do MARC 21 ao declarar que:

O formato MARC21 é um conjunto de códigos e designadores de conteúdo definido para a codificação de registros legíveis por máquina. É considerado um formato rico, completo e complexo, e sua estrutura se baseia na semântica estrutural do AACR2. Pode-se dizer que é o primeiro formato para comunicação de registros bibliográficos que descreve e identifica as características e especificidade do recurso informacional através de representação padronizada.

Na Tabela 1 estão os campos básicos do formato MARC 21.

Tabela 1 – Campos básicos do formato MARC 21.

Campos	Descrição
0XX	Informações de controle, números e códigos
1XX	Autoria (nome pessoal, entidade, evento)
2XX	Títulos, edição, imprensa
3XX	Descrição física
4XX	Série
5XX	Notas
6XX	Entradas de assunto
7XX	Entradas secundárias (nome pessoal, entidade, evento, título)
8XX	Entradas secundárias de série
9XX	Uso local

Fonte: Alves e Souza (2007)

Dublin Core

² A Divisão de Bibliotecas e Documentação da Pontifícia Universidade Católica do Rio de Janeiro (DBD/PUC-Rio) disponibiliza o acesso a todo formato MARC 21 no site: <http://www.dbd.puc-rio.br/MARC21/index.html/>

O Dublin Core é um padrão de metadados composto por apenas 15 campos além de qualificadores que são campos derivados dos 15 para melhor adaptação nas bibliotecas. Na [Tabela 2](#) estão os elementos do padrão Dublin Core.

Tabela 2 – Elementos metadados Dublin Core.

Elementos	Descrição
Título	Nome dado ao recurso
Criador	Entidade originalmente responsável pela criação do conteúdo do recurso
Assunto	Tema do conteúdo do recurso. Pode ser expresso em palavras-chaves e/ou Categoria. Recomenda-se o uso de vocabulários controlados
Descrição	Relato do conteúdo do recurso. Exemplos: texto livre, sumário e resumo
Publicador	Entidade responsável por tornar o recurso disponível
Colaborador	Entidade responsável pela contribuição intelectual ao conteúdo do recurso
Data	Data associada a um evento ou ciclo de vida do recurso
Tipo	Natureza ou gênero do conteúdo do recurso. Exemplos: texto, imagem, som, dados, software
Formato	Manifestação física ou digital do recurso. Exemplos: html, pdf, ppt, gif, xls
Identificador	Referência não-ambígua (localizador) para o recurso dentro de dado contexto
Fonte	Referência a um recurso do qual o presente é derivado
Idioma	Língua do conteúdo intelectual do recurso
Relação	Referência para um recurso relacionado
Cobertura	Extensão ou escopo do conteúdo do recurso; pode ser temporal e espacial
Direitos autorais	Informação sobre os direitos assegurados dentro e sobre o recurso

Fonte: [Alves e Souza \(2007\)](#)

Diferente do MARC que busca abarcar todos os dados possíveis de um documento, o Dublin Core busca simplificar o máximo o registro abarcando apenas os elementos suficientes para identificar o documento.

[Souza, Vendrusculo e Melo \(2000, p.94\)](#) afirmam que “as principais características do padrão DC são a simplicidade na descrição dos recursos, entendimento semântico universal (dos elementos), escopo internacional e extensibilidade (o que permite sua adaptação às necessidades adicionais de descrição)”. Com o objetivo de ser um formato simples no qual qualquer pessoa possa compreender, o Dublin Core busca simplificar o nome dos campos, dessa forma, cada metadado tem um nome que reflete diretamente o que quer dizer a descrição, por exemplo, onde se registra o título chama-se ‘título’ e onde se registra o autor chama-se ‘criador’.

A principal característica do formatado Dublin Core é: ser o padrão de interoperabilidade de metadados escolhido pelo protocolo Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH.

O protocolo OAI-PMH é um modelo de arquitetura de rede cliente-servidor que tem por objetivo regular tecnicamente como deve ocorrer o movimento dos metadados entre um provedor de dados e provedor de serviços, no contexto de um sistema federado de informações (MARTINS; FERREIRA, 2012, p.435).

Este protocolo ganha cada vez mais adeptos e já é considerado um dos principais protocolos de interoperabilidade entre repositório de conteúdos digitais (MARTINS; FERREIRA, 2012).

BibTeX

O BibTeX foi desenvolvido há por volta de 20 anos por Oren Patashnik. Ele foi concebido para ser utilizado em combinação com o processador de texto, L^AT_EX. A partir do BibTeX o L^AT_EX consegue gerar as citações no texto e a referência ao final do trabalho ³.

Neste formato os metadados partem do tipo de documento que se deseja catalogar, dessa forma, ao selecionar um tipo de documento como artigo, obtém-se apenas os metadados de artigo.

Assim como no Dublin Core, cada metadado possui um nome que se relaciona com o seu dado, então, onde se registra o título chama-se 'title' e onde se registra o autor chama-se 'author'.

O formato BibTeX abarca mais metadados que o Dublin Core, dessa forma, muitas adaptações que são feitas no Dublin Core por meio dos seus qualificadores não são necessárias serem feitas no BibTeX porque ele possui uma abrangência de descrição que abarca os campos necessários para cada tipo de documento.

Na Tabela 3 estão os metadados BibTeX para alguns tipos de documentos.

3.2 Tratamento temático

Quando o usuário não tem em mente um documento específico mas sim qualquer documento que seja pertinente a um assunto de interesse, ele estará diante de uma busca onde o sistema deve ter trabalhado o tratamento temático.

O tratamento temático, em bibliotecas, diz respeito ao assunto tratado no documento, ou seja, compreende a análise documentária como área

³ Informações encontradas no site oficial do BibTeX: <http://www.bibtex.org/>

Tabela 3 – Amostra Metadados BibTeX.

book	article	phdthesis
author	author	author
title	title	title
year	year	year
month	month	month
keywords	keywords	keywords
abstract	abstract	abstract
pages	pages	pages
address	address	address
note	note	note
number	number	type
volume	volume	school
publisher	journal	
editor		
edition		
series		

teórica e metodológica que abrange as atividades de classificação, elaboração de resumos, indexação e catalogação de assunto, considerando as diferentes finalidades de recuperação da informação (FUJITA et al., 2009, p.22).

É nesse fase onde se encontra a linguagem documentária que pode ser entendida como “um instrumento por meio da qual se realiza a mediação entre sistemas ou conjuntos informacionais e usuários” (LARA, 2004, p.233). Podem ser enquadradas dentro da linguagem documentária as classificações e os instrumentos utilizados na indexação como os tesauros e vocabulários de assuntos, por exemplo.

3.2.1 Classificação

O ato de classificar é natural ao humano, Siqueira (2010, p.38) afirma que “a necessidade de sistematizar o conhecimento é antiga, já que remonta à tentativa humana de representar e ordenar o mundo externo”. No cotidiano depara-se com a classificação, por exemplo, ao ir supermercado e ver as divisões entre produtos de limpeza e produtos de higiene pessoal ou quando em um ambiente social agrupa-se pessoas de acordo com algum critério como, sexo ou afinidade.

Souza (2009, p.16) define a classificação como “o processo de reunir coisas, ideias ou seres em grupos, de acordo com o seu grau de semelhança”. A classificação tem essa característica de reunir coisas que tenham algo em comum, porém “percebeu-se que os objetos eram agrupados não somente por semelhança, mas também a partir de uma ordem e distintos graus de diferenciação” (SIQUEIRA, 2010, p.39), isso porque características diferentes podem levar uma classe a se subdividir e estender a classificação. Nunes e Talamo (2009, p.34) defendem que a classificação “é a ordenação de um conjunto de seres

em pequenos agrupamentos, de acordo com características que os unem ou diferem de outros grupos”.

[Barbosa \(1969, p.43\)](#) divide as classificações em filosóficas e bibliográficas:

- a) Classificações filosóficas: são as puramente teóricas, constituindo agrupamentos dos conhecimentos humanos segundo o ponto-de-vista de seus idealizadores.
- b) Classificação bibliográficas é onde se encontra a CDD e a CDU que são as classificações voltadas para a organização de documentos.

[Souza \(2009, p.16\)](#) defende uma divisão mais recente e parecida. Partindo do ponto de vista da finalidade as classificações podem ser:

- a) Científica, quando sistematizam os fenômenos do mundo natural
- b) Documentárias, quando servem para organizar documentos, visando a recuperação da informação.

No contexto da Ciência da Informação as classificações de interesse são as documentárias que serão as utilizadas para apoiar a organização dos documentos.

Do ponto de vista de abrangência, [Souza \(2009, p.16\)](#) divide as classificações em:

- a) Especializadas, quando estão voltadas para um assunto determinado.
- b) Gerais, quando abrangem o universo do conhecimento.

[Nunes e Talamo \(2009, p.36\)](#), apontam que “classificação bibliográfica responde simultaneamente, a uma necessidade de organização interna das unidades de informação quanto de organização que vise à comunicação dos conteúdos armazenados com seus usuários”, essa colocação demonstra que a classificação bibliográfica deve atender simultaneamente interesses da organização e do usuário.

É possível o usuário ter acesso aos documentos do seu interesse por meio das relações comuns entre eles que a classificação bibliográfica demonstra, com isso ela apoia a pesquisa do usuário e coopera na construção de novos conhecimentos. [Siqueira \(2010, p.47\)](#) aponta que “a tentativa humana de estabelecer classes ou categorias mais do que uma ação supostamente de controle, é uma das formas mais hábeis de se representar uma realidade, trazendo subjacente o intuito de transformá-la em conhecimento”.

Dessa forma, dentre os objetivos finais da classificação documentária está o apoio ao acesso da informação e com isso a construção de conhecimento.

As principais classificações documentárias utilizadas no Brasil são CDD e a CDU, ambas são classificações que abrangem o universo do conhecimento.

3.2.1.1 CDD

A Classificação Decimal de Dewey - CDD foi desenvolvida em 1876 por Melvil Dewey. Partindo de pensamentos lógicos, Dewey dividiu o conhecimento humano em dez classes, de 0 a 9 sendo que as classes principais são escritas em centenas e se subdividem decimalmente (SOUZA, 2009). A CDD “atualmente é o sistema de classificação bibliográfica mais utilizada em todo o mundo, desde sua criação até os dias atuais passou por várias edições, sendo a de 2004 a mais atual, que corresponde a 22º edição” (ANDRADE; BRUNA; SALES, 2012, p.35).

O grande diferencial do CDD foi passar a utilizar os números na sequência decimal para definir as classes o que até então não tinha acontecido. Maine, em 1583, e Shurtleff, em 1856, apenas utilizaram a sequência decimal para numeração de prateleira e estantes (BARBOSA, 1969).

3.2.1.2 CDU

Baseados na CDD Paul Otlet e Henry La Fontaine criaram a Classificação Decimal Universal - CDU. A primeira edição, em francês, foi publicada em 1905 intitulada *Manuel du Répertoire de Bibliographie Universelle*, conhecida como Classificação de Bruxelas. Em 1931 foi publicada a segunda edição já com o nome que permanece até hoje, CDU (BARBOSA, 1969).

A CDU é dotada de uma estrutura inovativa, possuindo a habilidade de expressar, por símbolos, não apenas os assuntos simples, como também as relações entre os diversos assuntos. Sua estrutura é hierárquica, na qual o conhecimento é dividido em dez classes, podendo ser subdividido em partes lógicas, até o infinito (SOUZA, 2009).

O IBICT edita a versão portuguesa da CDU, desde, 1997 através da prerrogativa de ser o detentor da licença para edição em língua portuguesa, concedida pelo UDC Consortium – proprietário intelectual da CDU⁴. A CDU ter a versão portuguesa é um diferencial positivo com relação à CDD que não possui tradução para este idioma (ANDRADE; BRUNA; SALES, 2012).

A Tabela 4 mostra uma relação entre as 10 principais classes da CDD e da CDU.

As duas classificações possuem uma divisão de classes bem parecida a grande diferença é que a CDU passou a classe 400 da CDD, Línguas para a sua classe 8, 'Linguística. Literatura' e deixou a classe 4 vaga já antecipando o lugar para o surgimento de uma possível nova área de conhecimento no futuro.

⁴ Informações encontradas no site oficial do IBICT: <http://www.ibict.br/publicacoes-e-institucionais/classificacao-decimal-universal-cdu>

Tabela 4 – Principais classes da CDD e CDU.

CDD	CDU
000 Generalidades	0 Generalidades
100 Filosofia e disciplinas relacionadas	1 Filosofia
200 Religião	2 Religião
300 Ciências Sociais	3 Ciências Sociais
400 Línguas	4 Vaga
500 Ciências Puras	5 Ciências Puras
600 Tecnologia (Ciências Aplicadas)	6 Ciências Aplicadas
700 Artes, Recreação e Artes Cênicas	7 Artes. Recreação. Diversão. Esportes
800 Literatura (Belas Letras)	8 Linguística.Literatura
900 Geografia. Biografia. História	9 História, Geografia. Biografias

3.2.2 Indexação

Como dito anteriormente, a indexação faz parte do tratamento temático do documento juntamente com a classificação. A relevância desse tratamento fica clara na afirmação:

[...]a maior parte das demandas da informação submetidas pelos usuários apresenta um enfoque de natureza temática e não é habitual que conheçam de forma precisa nem a autoria nem os títulos das obras desejadas (FUJITA; LACRUZ; DÍAZ, 2012, p.96).

Possibilitar que o usuário encontre documentos apenas pelo assunto faz toda a diferença na prestação dos serviços de informação, como dito pelo autores, o usuário geralmente não possui um documento em mente tendo apenas os assuntos como ponto de partida da busca.

Para Lancaster (1993) a indexação de assuntos envolve duas etapas principais, a análise conceitual e a tradução.

Na análise conceitual será feita a decisão sobre o que trata o documento e isso envolve todo o trabalho inicial de buscar os assuntos daquele documentos tendo em mente sempre o usuário para que essa análise seja feita de forma a privilegiar os assuntos que possam ser de seu interesse, pois “a mesma publicação pode ser indexada de forma bastante diferente, se os grupos de usuários estiverem interessados nesse documento por diferentes razões” (LANCASTER, 1993, p.8).

A segunda etapa, tradução, envolve a conversão da análise conceitual para um determinado conjunto de termos. Essa conversão da análise conceitual em tradução deve ser feita com muita cautela, pois compreender o assunto de um documento é diferente de

dar nomes a partir dessa compreensão, isso porque o bibliotecário deve ter cuidado de não escolher um termo que não seja pertinente à análise feita.

A definição dos termos pode partir de uma indexação por extração (indexação derivada) ou indexação por atribuição. Na indexação por extração, os termos são definidos a partir de palavras que ocorrem no próprio documento. A indexação por atribuição já envolvem as linguagens documentárias pois a atribuição de termos a um documento é feita a partir de uma fonte que não é o próprio documento (LANCASTER, 1993).

Por partir de uma linguagem controlada, a indexação por atribuição consegue evitar problemas como incluir novos termos no catálogos que já possuem outras nomeações com o mesmo conceito. Esse problema é um dos exemplos dentre vários que podem ocorrer quando se adota a indexação por extração, por isso que adotar uma linguagem documentária é aconselhável.

As bases conceituais da Linguística Documentária é formada por conceitos de três áreas:

[...]da Lingüística, o princípio estrutural de organização da linguagem; da Lógica, a identificação das formas de raciocínio e de organização de conjuntos; da Terminologia, a modelagem do conceito e dos sistemas de conceitos, além das referências concretas para a interpretação dos termos por meio dos glossários e dicionários terminológicos especializados, que são seus produtos (LARA, 2004, p.232).

Tendo em vista as bases que as linguagens documentárias necessitam para serem formadas, é possível perceber que construir um tesauro, por exemplo, não é uma tarefa fácil e requer muito trabalho e empenho.

Além de possuir uma linguagem documentária, para ser feita uma boa indexação também é necessária uma política de indexação. O indexador irá executá-la a partir das diretrizes presentes nesta política, com isso, a padronização no processo será mais eficiente (FUJITA et al., 2009).

Parte III

Resultados

4 Introdução

Este relato de caso busca mostrar toda a trajetória percorrida pela equipe do projeto biblioteca digital do CPAI privilegiando os aspectos de organização da informação tratados na revisão de literatura. Para contextualizar o relato de caso, no início são apresentados o centro de pesquisa e o surgimento do projeto.

Após a contextualização, são apresentados os requisitos levantados no início do projeto que foram a base para orientar as buscas por um ferramenta para a biblioteca digital. Na sequência as experiências com a primeira e a segunda ferramenta são apresentadas de forma a mostrar os pontos referentes à utilização e testes com essas ferramentas e questões relacionadas aos metadados.

Por fim os detalhes das experiências com a classificação e indexação são relatados.

5 Relato do caso

5.1 Sobre o CPAI

Fundado em 10 de maio de 2007, o Centro de Pesquisa em Arquitetura da Informação – CPAI, vinculado à Faculdade de Ciência da Informação - FCI, da Universidade de Brasília, é constituído por um quadro de pós-doutores, doutores, mestres, estudantes de doutorado, de mestrado e de graduação envolvidos em pesquisas sobre a estruturação e uso de Arquitetura da Informação. O foco do centro está na inovação em Arquitetura da Informação na Sociedade Sustentável tendo como missão promover pesquisa, desenvolvimento e inovação em Arquitetura da Informação de forma transdisciplinar para oferecer soluções de sustentabilidade à Sociedade. A visão do CPAI é ser um centro de vanguarda e excelência em epistemologia, ciência e práxis na Arquitetura da Informação, consolidando a Escola de Brasília como referência para a Sociedade da Informação tendo como valores a ética, a responsabilidade social – reciprocidade social (*accountability*) e a excelência.

Para atender às suas finalidades o CPAI está estruturado em um modelo organizacional matricial dividido em núcleos com o objetivo de direcionar esforços para áreas bem definidas.

Essa estrutura é representada na figura [Figura 1](#):

Figura 1 – Estrutura do CPAI



Fonte: <http://cpai.unb.br/pt/estrutura-organizacional>

Os serviços da biblioteca digital terão como foco o Núcleo de Pesquisa que é a

unidade responsável pela coordenação dos projetos de pesquisa cuja temática é de natureza científica e relacionada aos fundamentos da Arquitetura da Informação ou de áreas de interesse desta, além de que este núcleo mantém um programa educacional de caráter formativo com programas de graduação e pós-graduação em Arquitetura da Informação¹.

5.2 Surgimento do projeto

Grande parte do quadro de pessoal do CPAI está envolvido com o ambiente acadêmico, escrevendo suas teses, dissertações, monografias ou artigos científicos. Com quase 10 anos de fundação, muitas pesquisas já foram concluídas com o apoio do CPAI que deram origem a produção de documentos como os mencionado anteriormente. Tornou-se necessária uma solução que controlasse e facilitasse o acesso a estes. Reunindo essa demanda de publicações vinculadas ao centro de pesquisa e de outros documentos que possam ser de interesse dos pesquisadores do centro, surgiu o projeto Biblioteca Digital do CPAI - BD CPAI.

O projeto foi definido com o nome de BD CPAI, mas porque não foi chamado de Biblioteca Virtual do CPAI ou Biblioteca Eletrônica do CPAI? Como visto na revisão de literatura o conceito de biblioteca virtual se relaciona com a realidade virtual como defendido por Marchiori (2009) e com os catálogos virtuais ou a própria *World Wide Web* mencionados por Tammaro e Salarelli (2008). Como a BD CPAI não tem a intensão de explorar tecnologias da realidade virtual e foi definida para ser uma biblioteca fechada não se relacionando com a grandeza da *World Wide Web* ou a dos catálogos virtuais então não cabia a utilização da expressão ‘biblioteca virtual’. A biblioteca eletrônica possui como a sua principal característica a utilização de recursos eletrônicos assim como defendido por Tammaro e Salarelli (2008), Marchiori (2009) e Martins (2002), dessa forma a BD CPAI poderia se chamar de Biblioteca Eletrônica do CPAI, porém os conceitos de bibliotecas digitais reúnem mais características como as apontadas por Sayão (2009) e Tammaro e Salarelli (2008) que se adequam melhor à proposta do projeto do CPAI e por isso foi escolhida a expressão ‘biblioteca digital’.

Como levantado na revisão de literatura, para ser uma biblioteca digital é necessário ter um público alvo que se tornará os usuários da biblioteca. Este projeto já tem delimitado os usuários desde o seu início pois, o surgimento do projeto partiu da necessidade informacional deles. O conteúdo da biblioteca também é claro sendo os documentos que servirão de apoio para as pesquisas do centro e os próprios documentos de publicação vinculadas ao CPAI. Desta forma, o acervo é constituído de artigos científicos, teses, dissertações, monografias e livros. Como o centro de pesquisa é voltado para uma área específica, Arquitetura da Informação, os trabalhos produzidos com o apoio do centro

¹ As informações desta sessão foram encontradas no portal do CPAI: <http://cpai.unb.br/home>

geralmente possuem pontos em comum o que acaba fazendo com que o trabalho de um pesquisador muitas vezes colabora com o do outro. Dessa forma é fundamental que todos os trabalhos envolvidos com o centro estejam na biblioteca não só por ser de interesse do centro reunir essas publicações mas porque certamente será de proveito para o trabalho dos pesquisadores.

Os usuários da biblioteca possuem relação estreita com as tecnologias sendo acostumados a fazerem buscas por informações na *internet* com isso é muito interessante o acesso aos documentos de casa por meio de uma biblioteca digital, dessa forma, a BD CPAI será disponibilizada em rede para que esses usuário possam ter acesso à ela de onde estiverem, ainda mais que, estar disponível em rede é uma das características das bibliotecas digitais.

5.3 Levantamento de requisitos

Para iniciar o projeto, foi necessário levantar os requisitos a serem atendidos pela ferramenta que irá ser utilizada na BD CPAI. Este levantamento é fundamental para que a equipe tenha clareza do que ela realmente espera da ferramenta e conseqüentemente da biblioteca digital. Foram definidos os requisitos dividindo-os em: requisitos mínimos e requisitos desejáveis.

Requisitos mínimos

A [Tabela 5](#) mostra os requisitos mínimos juntamente com a sua justificativa.

Requisitos desejáveis

A [Tabela 6](#) mostra os requisitos desejáveis juntamente com a sua justificativa.

5.4 Experiência com a primeira ferramenta

Com os requisitos levantados, iniciou-se uma fase de busca pelas ferramentas utilizadas nas bibliotecas digitais de todo o mundo.

Uma ferramenta muito utilizada mundialmente é o DSpace, que possui uma comunidade de usuários no Brasil com mais de 60 adeptos. Por meio de entrevistas, via telefone, com os responsáveis das bibliotecas digitais de instituições em todo Distrito Federal e ficou confirmado que o DSpace é a ferramenta mais utilizada no DF.

Apesar de ter levantado outras ferramentas como o Eprints e o Greestone, o projeto BD CPAI acabou adotando a ferramenta DSpace porque atendia a todos os requisitos levantados e possuía uma comunidade de usuários próximos ao centro de pesquisa o que

Tabela 5 – Requisitos mínimos e justificativa.

Requisitos mínimos	Justificativa
Plataforma <i>web</i>	para que a biblioteca tenha o acesso em rede
<i>Open source</i>	para que seja possível modificar o código da ferramenta caso seja necessário
Livre	para que não seja necessário pagar pela ferramenta
Inclusão e exclusão de documentos e links	para disponibilizar o documento e fazer referências a links de outros sites
Tratamento descritivo dos documentos	para ter formas de catalogar o documento
Tratamento temático dos documentos	para ter formas de indexar e classificar o documento
Busca por autor	para que a ferramenta possibilite fazer buscas pelo autor
Busca por título	para que a ferramenta possibilite fazer buscas pelo título
Busca por assunto	para que a ferramenta possibilite fazer buscas pelo assunto
Interoperabilidade	para que o sistema converse com outros sistemas de bibliotecas
Controle de acesso	para distinguir ações de gerenciamento e consulta

Tabela 6 – Requisitos desejáveis e justificativa.

Requisitos desejáveis	Justificativa
Busca textual	para ser possível fazer a busca pelas palavras do texto
Integração com o Liferay	para integrar com o portal Liferay que é a utilizada no portal do CPAI
Controle de acesso por item	para controlar o acesso aos documentos

facilita caso seja necessário buscar por experiências de outras bibliotecas. Outra característica que influenciou na escolha do DSpace é o suporte que o Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT presta para as bibliotecas que utilizam esta ferramenta.

Decidido que a ferramenta a ser adotada seria o DSpace, esta foi instalada no servidor do centro de pesquisa e após isso a equipe do projeto da BD CPAI iniciou os primeiros testes.

O primeiro ponto observado foi a estruturação do DSpace que se organiza em: comunidades, sub comunidades e coleções. Somente nas coleções é possível adicionar documentos, então as comunidades podem se sub dividir em coleções ou em sub comunidade e estas serem sub divididas em coleções. Por exemplo, se houver a seguinte divisão:

1. Ciência da Informação

1.1. Biblioteconomia

1.1.1. Classificação Bibliográfica

No Dspace a ‘Ciência da Informação’ seria uma comunidade, a ‘Biblioteconomia’ uma sub comunidade e ‘Classificação Bibliográfica’ uma coleção. Só seria possível incluir documentos dentro de ‘Classificação Bibliográfica’ caso houvesse um documento que se queira incluir em ‘Biblioteconomia’ ou ‘Ciência da Informação’ não seria possível. Outro ponto é a necessidade de classificar um mesmo documento em mais de uma classe, no Dspace é possível incluir um documento em uma coleção e espelhar ele em outras coleções porém não é possível espelhar nas comunidade e sub comunidades. Como a BD CPAI trabalha com uma classificação² onde um mesmo documento pode ser de mais de uma classe e pode estar presente em qualquer nível da hierarquia da classificação, essa estruturação não se adequou bem a demanda. Concluiu-se que esse esquema de divisão adotado pelo Dspace não serve para receber uma classificação da BD CPAI.

Outro ponto observado foram os metadados, o Dspace utiliza o padrão de metadados Dublin Core. Como explanado na revisão de literatura, ele é um formato simples que pode ser expandido através dos seus qualificadores. O formato Dublin Core foi feito para aceitar adaptações, e no Dspace é possível observar adaptações feitas neste padrão de metadados.

A [Tabela 7](#) mostra uma relação entre os metadados do Dublin Core sem adaptações e os do Dublin Core do Dspace na forma com vem o software sem modificações.

Tabela 7 – Metadados Dublin Core e Dspace.

Dublin Core	Dspace
Título	Title
Criador	Authors
Assunto	Subject
Descrição	Description
Publicador	Publisher
Colaborador	Sponsors
Data	Date
Tipo	Type
Formato	Format
Identificador	Identifier
Fonte	
Idioma	Language
Relação	
Cobertura	
Direitos autorais	License
	Citation
	Series/Report No.
	Abstract

A partir disso observa-se que a interface do Dspace não oferece os metadados ‘fonte’, ‘relação’ e ‘cobertura’ para serem preenchidos e que foram adicionados dois novos metadados: ‘citation’, ‘series/report no.’ e ‘abstract’

² Esta classificação será abordada com detalhes na [seção 5.6](#)

Quando iniciou-se os testes de inserção de documentos observou-se a necessidade de clareza na hora de preencher os campos. Por exemplo, nos documentos acadêmicos como teses e dissertações os dados importantes para descrição são: título, autor, escola, ano, assunto e resumo. No Dspace, claramente existem os metadados para título, autor, ano, assunto e resumo, porém não existe metadado claro para ‘escola’. Poderia-se fazer uma adaptação e inserir a ‘escola’ em ‘publisher’ mas havendo um campo claro para escola independente da pessoas que estiver catalogando o documento a o dado sobre a escola será preenchido da mesma forma.

O que mudou os rumos do projeto e acabou levando à mudança de ferramenta foi uma demanda que até então não tinha sido tão clara e evidente: adoção de metadados BibTex. Como dito anteriormente, dentre os usuários do CPAI encontram-se estudantes de doutorado, de mestrado e de graduação. Com o intuito de padronizar os trabalhos acadêmicos vinculados ao centro, o CPAI adotou como processador de texto o L^AT_EX. Dessa forma, todos os trabalhos envolvidos com o centro tiveram que adotar o L^AT_EX. A partir dessa decisão, ter o formato de metadados BibTex dos documentos passou a ser um grande interesse para todos os usuários.

O Dspace não exporta para BibTex então iniciou-se um trabalho para ver a possibilidade de transformar as descrições feitas pelo Dspace em BibTex. Como o BibTex é um formato mais extenso e detalhado que o Dublin Core adotado pelo Dspace percebeu-se que essa adaptação daria muito trabalho. Na [Tabela 8](#), uma relação entre a amostra de metadados BibTex da [Tabela 3](#) e os metadados do Dublin Core do Dspace da [Tabela 7](#) para mostrar as diferenças que aparecem.

Tabela 8 – Metadados BibTeX e Dspace.

BibTeX	Dspace
title	Title
author	Authors
keywords	Subject
Series/ number	Series/Report No.
publisher	Publisher
Year/ month	Date
type	Type
abstract	Abstract
	Description
	Sponsors
	format
	Identifier
	Language
	License
	Citation
pages	
address	
note	
school	
editor	
edition	
type(para trabalho acadêmico)	
volume	

Juntando a inadequação da estrutura de comunidade e coleções e a insatisfação com os metadados, decidiu-se mudar de ferramenta.

5.5 Experiência com a segunda ferramenta

Na escolha da segunda ferramenta o principal requisito passou a ser a utilização do metadado BibTeX.

Em buscas direcionadas pela *internet* encontrou-se uma comparação entre softwares de gerenciamento de referencia que nortearam as buscas por uma nova ferramenta³.

A partir dos softwares presentes nessa comparação houve um trabalho de refinamento escolhendo apenas os que atendem os seguintes requisitos:

- a) Plataforma *web*;
- b) *Open source*;
- c) Livre;
- d) Importa BibTeX;
- e) Exporta BibTeX;

³ Essa comparação está em uma página do Wikipedia: http://en.wikipedia.org/wiki/Comparison_of_reference_management_software

f) Rodar nos principais sistemas: Windows, Mac OS, Linux e Unix.

Ao final deste refinamento chegou-se na ferramenta Aigaion que possui como forma de estradas dos dados o padrão BibTeX. O formato BibTeX foi visto pela equipe da BD CPAI como um formato mais claro e completo comparado ao formato de metadados utilizado pela ferramenta Dspace.

A forma de classificar os documentos se mostrou mais simples e clara comparada a da ferramenta anterior. Trabalhando com um sistema de tópico e sub tópicos é possível adicionar camadas de sub classes o quanto for necessário. No Aigaion não é necessário escolher uma classe para depois incluir um documento dentro dela como ocorre no Dspace. Na nova ferramenta a classificação é feita depois da descrição do documento em que é apresentada a classificação e é possível escolher em quais classes aquele documento sera incluído, dessa forma, a classificação pode ser feita em qualquer nível e em quantas classes forem necessárias.

Na experiência com a segunda ferramenta outras necessidades mais pontuais foram surgindo, por exemplo, a inclusão do metadado orientador e a possibilidade de fazer buscas por este campo. A inclusão do metadado orientador foi feita, porém até o momento a demanda de busca com base neste campo ainda não é possível.

Com esse exemplo do campo de orientador, ficou claro para a equipe que os metadados de entrada de dados nem sempre são os metadados utilizados na busca e que necessariamente para se fazer a busca por um campo, este deve ter sido preenchido na entrada.

A adoção dessa nova ferramenta causou discussões sobre um dos requisitos mínimos: interoperabilidade. Na ferramenta anterior, Dspace, essa interoperabilidade era atendida devido a utilização do padrão de metadados Dublin Core que, como visto na revisão de literatura, é o padrão adotado pelo protocolo OAI-PMH que possibilita a comunicação entre os sistemas. No Aigaion o padrão é o BibTeX, sendo assim, essa ferramenta não possui a interoperabilidade já que não utiliza o padrão de metadados definido mundialmente pelo protocolo OAI-PMH. Porém isso põe em pauta uma questão: é mais importante atender da melhor forma os usuários ou uma demanda de interoperabilidade? Adotando o padrão BibTeX os usuários são atendidos da melhor forma pois este é o padrão que eles precisam para apoiar o seus trabalhos acadêmicos, porém para que a biblioteca tenha interoperabilidade ela precisa conversar com o protocolo OAI-PMH. Esta questão de interoperabilidade deve ser melhor trabalhada para que a BD CPAI não fique prejudicada caso tenha que se comunicar com outra biblioteca.

5.6 Experiência com a classificação

Falar de classificação em uma biblioteca digital aborda questões que antes não ocorriam com a biblioteca tradicional. Pensando em uma exemplo como a Biblioteca Central da Universidade de Brasília - BCE-UnB: o usuário chega na biblioteca, procura algum documento no catálogo eletrônico, anota o número de chamada, que é a classificação, e em seguida com esse número procura o documento na estante. Este exemplo mostra uma das principais funções que as classificações nas bibliotecas tradicionais possuem, definir um local para o documento na estante de forma que ele possa ser encontrado posteriormente. Em uma biblioteca digital, essa ideia de “achar o documento na estante” não existe. Ao entrar no sistema de busca de uma biblioteca digital e encontrar os dados do documento que procura, o usuário já encontra também o acesso a este em formato digital.

Mas a classificação não é só para “achar o documento na estante”. Ela coloca documentos com assuntos relacionados próximos e assim, o usuário ao procurar algo específico pode encontrar outros documentos interessantes próximo do que tinha em mente. Isso adiciona mais possibilidades de o usuário encontrar mais informações que o interessa.

No início da tentativa de definir uma classificação para a BD CPAI foi questionado a sua necessidade ao se ter em mente apenas a questão de “achar o documento na estante” mas quando ficou perceptível a importância e eficácia da classificação ao mostrar de forma simples e clara os documentos relacionados não houve dúvida de sua utilização.

A experiência da elaboração da classificação da biblioteca foi enriquecedora para os integrantes da equipe da BD CPAI, pois foi feita com o apoio direto do diretor do centro de pesquisa reunindo com equipe para elaborar a classificação mais adequada para o centro.

De início decidiu-se usar a CDU por ser a classificação que a equipe mais tinha afinidade daí a partir das 10 principais classes levantou-se as que são de interesse do centro. A seguir na [Tabela 9](#) as classes escolhidas para o projeto.

Tabela 9 – Classes da CDU para o projeto.

0	Generalidades
1	Filosofia
3	Ciências Sociais
5	Ciências Puras
6	Ciências Aplicadas
8	Linguística.Literatura

Com as grandes classes definidas iniciou-se o trabalho com as subclasses. A ideia de trabalho foi a mesma das grandes classes, levantou-se as subclasses da CDU e a partir disso delimitou as que fariam parte da classificação do centro. Até então nada foi adicionado na classificação da CDU, apenas retirado, o que poderia levar o projeto a apenas incluir

a CDU dentro da biblioteca digital e questão encerrada. Mas a partir do nível seguinte às subclasses começaram as inclusões de classes diferentes das existentes na CDU.

Em reuniões com o diretor novas subclasses foram surgindo e adaptações foram feitas na CDU até chegar na classificação final.

Essa experiência de levantar uma classificação voltada para as necessidades do centro de pesquisa foi importante para que a equipe da biblioteca digital compreendesse melhor o que entra no escopo de assuntos de interesse do centro de pesquisa. Como o projeto BD-CPAI ainda era recente a equipe não tinha muito conhecimento do seu público alvo e o trabalho com o levantamento da classificação muitos assuntos que antes pareciam não ter relação com os interesses do CPAI se mostraram claramente dentro do escopo.

A classificação não é algo rígido e por isso está aberta a novas adaptações que sejam convenientes. Um exemplo é a inserção da classe 7 ‘Artes. Recreação. Diversão. Esportes’. A equipe da biblioteca digital começou a se deparar com muitos documentos da área de Design, como ainda não havia uma classe para essa área começou-se a se discutir a inserção de uma nova classe. Como os documentos tinham origem dos acervos particulares dos pesquisadores era claro o interesse desse tema para os usuários da biblioteca e com isso ficou decidido incluir uma nova classe. Decidido que a nova classe entraria, a equipe do projeto juntamente com um dos pesquisadores do centro e da área de design definiram as divisões para esta nova classe e ela foi incluída na classificação juntamente com suas subdivisões. A classificação final da BD CPAI pode ser vista no apêndice A.

5.7 Experiência com a indexação

Como visto na revisão de literatura, a indexação é o processo onde os assuntos do documento são levantados e definidos sendo dividido em análise conceitual e tradução. No projeto BD CPAI a análise conceitual é feita tendo em mente os assuntos de interesse do centro de pesquisa visto que o centro possui um rol de assuntos específicos de seu interesse. Feita essa análise parte-se para a fase de tradução. Essa fase pode ser feita tanto por extração como por atribuição. Para que haja uma indexação por atribuição é necessário que exista uma linguagem documentária.

Como visto na revisão de literatura, a indexação mais recomendada é a indexação por atribuição, isso porque é apoiada por uma linguagem documentária que evita erros na indexação. Porém, a construção de uma linguagem documentária requer muito trabalho e o envolvimento de conceitos de linguística, lógica e terminologia. Tendo em vista que a construção da linguagem documentária levaria um tempo e para que os documentos não fossem incluídos na biblioteca sem uma indexação, a equipe decidiu por adotar a indexação por extração.

A equipe foi percebendo na prática as dificuldades encontradas quando se adota uma indexação por extração. De início, foi definido que as palavras chave devem sempre serem feitas em português independente da língua do documento. Isso porque é certo que todos os usuário da biblioteca digital conhecem a língua portuguesa. Porém, percebeu-se que existem assuntos que são mais conhecidos na língua inglesa e isso mostrou a necessidade de entrar tanto com a palavra em língua portuguesa como em língua inglesa. Uma linguagem documentaria poderia fazer relações de equivalência e relacionar esses termos que possuem o mesmo conceito porém estão em línguas diferentes para que o usuário possa usar qualquer uma das formas e encontrar todos os documentos com aquele mesmo conceito. Este exemplo é um dentre outros que mostraram para a equipe a necessidade de adotar uma linguagem documentária.

Considerações finais e trabalhos futuros

Relembrando as palavras de Cunha (2008, p.6) sobre os projetos de bibliotecas digitais “os atuais gerentes deste tipo de projeto estão aprendendo com a prática!”. O projeto BD CPAI é um exemplo desses projetos que tiveram que aprender com a prática e que dentre várias decisões tomadas umas deram certo e outras demandaram mais trabalho e novas decisões.

As conclusões deste trabalho, apontam claramente para trabalhos futuros, dessa forma serão apresentadas os trabalhos futuros juntamente com suas justificativas:

- a) Estudar a questão de interoperabilidade: a ferramenta Aigaion atende bem as demandas da BD CPAI, porém a questão de interoperabilidade ficou fragilizada. Para que a BD CPAI não fique prejudicada caso precise conversar com outras bibliotecas é necessário que se busque ferramentas que façam a conversão entre os dois formatos ou trabalhar com a possibilidade de adotar uma nova ferramenta que importe e exporte tanto para BibTeX como para Dublin Core.
- b) Estabelecer detalhadamente políticas de catalogação: no projeto BD CPAI o processo de catalogação ficou com o foco nos metadados, porém, as regras para o preenchimento desses campos são muito importantes para que este processo seja feito da melhor forma. O código AACR2R deve ser adotado para que este processo melhore os seus serviços.
- c) Elaborar uma linguagem documentária e uma política de indexação: a indexação de extração é vulnerável à muitos erros e para que a BD CPAI melhore as suas indexações é necessário que se adote uma política de indexação juntamente com uma linguagem documentária. Atualmente o projeto BD CPAI está com um novo subprojeto que visa a elaboração de uma linguagem documentária. Até o momento percebe-se que as palavras chave já levantadas ajudarão muito no início do projeto para que a nova equipe tenha uma visão dos assuntos dos documentos levantados até então.
- d) Fazer um Estudo de usuário: com o intuito de verificar a utilização da biblioteca e buscas formas de melhorar os serviços.

As ferramentas de bibliotecas digitais existem para atender as demandas de organização e disponibilização de informações. Estas demandas são levantadas à partir de conhecimentos provindos da área de ciência da informação. Dessa forma, é importante que as questões tecnológicas que envolvem essas bibliotecas sem paredes não sobreponham

às da área de ciência da informação, para que, a equipe não perca o foco dos aprendizados fundamentais desta área devido a uma maior preocupação com a tecnologia.

Referências

- ALVES, M. das D. R.; SOUZA, M. I. F. Estudo de correspondência de elementos metadados: Dublin core e marc 21. *Revista Digital de Biblioteconomia e Ciência da Informação*, 2007. Campinas, v. 4, n. 2, p. 20–38, 2007. Citado 4 vezes nas páginas [39](#), [40](#), [41](#) e [42](#).
- ANDRADE, L. V. de; BRUNA, D.; SALES, W. N. de. Classificação: uma análise comparativa entre a classificação decimal universal - cdu e a classificação decimal de deey - cdd. *Biblos: Revista do Instituto de Ciências Humanas e da Informação*, 2012. v. 25, n. 2, p. 31–42, 2012. Citado na página [46](#).
- BARBOSA, A. P. *Teoria e prática dos sistemas de classificação bibliográfica*. Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação, 1969. Citado 2 vezes nas páginas [45](#) e [46](#).
- CAMPELLO, B. *Introdução ao controle bibliográfico*. Brasília: Briquet de Lemos/Livros, 2006. Citado na página [38](#).
- CERVO, A. L.; BERVIAN, P. A.; SILVA, R. da. *Metodologia científica*. 6. ed. São Paulo: Pearson Prentice Hall, 2007. Citado na página [24](#).
- COZBY, P. C. *Métodos de Pesquisa em Ciências do Comportamento*. São Paulo: Atlas, 2003. Citado na página [24](#).
- CUNHA, M. B. da. Desafios na construção de uma biblioteca digital. *Ciência da Informação*, 1999. v. 28, n. 3, p. 257–268, 1999. Citado 2 vezes nas páginas [19](#) e [33](#).
- CUNHA, M. bastos da. Das bibliotecas convencionais às digitais: diferenças e convergências. *Perspectiva em Ciência da informação*, 2008. v. 13, n. 1, p. 2–17, 2008. Citado 3 vezes nas páginas [23](#), [31](#) e [65](#).
- FUJITA, M. S. L. et al. *A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias*. São Paulo: Cultura Acadêmica, 2009. Citado 3 vezes nas páginas [37](#), [44](#) e [48](#).
- FUJITA, M. S. L.; LACRUZ, M. del C. A.; DÍAZ, R. G. A situação atual da indexação nas tarefas bibliotecárias. *Perspectivas em Ciência da Informação*, 2012. São Paulo, v. 17, n. 1, p. 94–109, 2012. Citado na página [47](#).
- FUSCO, E. *Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais*. São Paulo: Cultura Acadêmica, 2011. Citado 4 vezes nas páginas [38](#), [39](#), [40](#) e [41](#).
- LANCASTER, W. F. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos/Livros, 1993. Citado 2 vezes nas páginas [47](#) e [48](#).
- LANDONI, M.; CATENAZZI, N.; GIBB, F. Hyper-books and visual-books in an electronic library. *The Electronic Library*, 1993. v. 11, n. 3, p. 175–176, 1993. Citado na página [31](#).

- LARA, M. L. G. de. Linguagens documentárias e terminologia. *Trasinformação*, 2004. Campinas, v. 16, n. 5, p. 231–240, 2004. Citado 2 vezes nas páginas 44 e 48.
- LEVACOV, M. Bibliotecas virtuais:(r)evolução? *Ciência da Informação*, 1997. v. 26, n. 2, 1997. Citado 2 vezes nas páginas 31 e 32.
- MARCHIORI, P. Z. “ciberteca” ou biblioteca virtual: uma perspectiva de gerenciamento de recursos de informação. *Ciência da Informação*, 2009. v. 38, n. 2, 2009. Citado 4 vezes nas páginas 31, 32, 33 e 54.
- MARCONI, M. de A.; LAKATOS, E. M. *Metodologia científica*. 6. ed. São Paulo: Atlas, 2011. Citado na página 25.
- MARTINS, D. L.; FERREIRA, S. M. S. P. Protocolo oai-pmh e sistemas federados de informação: fundamentos de arquitetura da informação para análise de dados do portal de produção científica da área de ciência da computação univerciencia.org. *Liinc em Revista*, 2012. Rio de Janeiro, v. 8, n. 2, p. 431–447, 2012. Citado na página 43.
- MARTINS, R. D. Perspecticas para uma biblioteca no futuro: utopia ou realidade. *Informação e Sociedade: Estudos*, 2002. v. 12, n. 1, 2002. Citado 2 vezes nas páginas 33 e 54.
- MEY, E. S. A.; SILVEIRA, N. C. *Catalogação no plural*. Brasília, DF: Briquet de Lemos/Livros, 2009. Citado 4 vezes nas páginas 37, 38, 39 e 41.
- NAVES, M. M. L.; KURAMOTO, H. *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos/Livros, 2006. Citado 4 vezes nas páginas 31, 32, 34 e 37.
- NUNES, L.; TALAMO, M. de F. G. M. Da filosofia da classificação à classificação bibliográfica. *Revista Digital de Biblioteconomia e Ciência da Informação*, 2009. Campinas, v. 7, n. 1, p. 30–48, 2009. Citado 2 vezes nas páginas 44 e 45.
- SAYÃO, L. F. Afinal, o que é biblioteca digital? *Revista USP*, 2009. n. 80, p. 6–17, 2009. Citado 4 vezes nas páginas 32, 34, 35 e 54.
- SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 2010. Florianópolis, v. 15, n. 30, p. 1–13, 2010. Citado na página 40.
- SIQUEIRA, J. C. O conceito classificação: uma abordagem história e epistemológica. *Revista Brasileira de Biblioteconomia e Documentação*, 2010. São Paulo, v. 6, n. 1, p. 37–49, 2010. Citado 2 vezes nas páginas 44 e 45.
- SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão dublin core. *Ciência da Informação*, 2000. Brasília, v. 29, n. 1, p. 93–102, 2000. Citado na página 42.
- SOUZA, S. de. *CDU: como entender e utilizar a 2ª edição-padrão internacional em língua portuguesa*. Brasília: Thesaurus, 2009. Citado 3 vezes nas páginas 44, 45 e 46.
- TAMMARO, A. M.; SALARELLI, A. *A biblioteca digital*. Brasília: Briquet de Lemos/Livros, 2008. Citado 4 vezes nas páginas 23, 33, 34 e 54.

Apêndices

APÊNDICE A – Classificação BD CPAI

A classificação da BD CPAI está apresentada a seguir da mesma forma que aparece na biblioteca. Os números que aparecem após cada classe são referentes à quantidade de documentos que ela possui.

Figura 2 – Classe 0

- ▣ Classe 0 1/659
 - ▣ Ciência da Informação 1/653
 - ▣ Arquitetura da Informação 1/371
 - ▣ Ciência 1/156
 - Cadeia Normativa da Segurança da Informação – CNSI 0/6
 - Cadeia Normativa Organizacional 0/0
 - Método de Arquitetura da Informação Aplicada – MAIA 0/2
 - Novos modelos de Aplicação para Lógicas Modais 0/0
 - Teoria em Programação Lógica Clássica e Lógicas Não- Clássicas 0/0
 - Teoria Geral da Arquitetura da Informação 1/9
 - Teorias e métodos sobre planejamento estratégico organizacional 1/32
 - Teorias sobre Arquitetura de Sistemas de Informação 0/104
 - ▣ Epistemologia 0/42
 - A teoria do conhecimento: fenomenologia 0/2
 - Estudos sobre a adoção da fenomenologia 0/0
 - Fundamentos de Arquitetura da Informação 0/17
 - Lógicas Clássica e Não-Clássicas 0/0
 - Natureza do Dado, Informação e Conhecimento 0/22
 - Teorias em Programação em Lógica 0/1
 - ▣ Práxis 0/215
 - Arquitetura da Garantia da Informação 0/5
 - Arquitetura da Geogestão 0/4
 - Arquitetura da Informação Escolhas Sociais 0/0
 - Arquitetura da Informação Estratégica 0/50
 - Arquitetura da Informação Organizacional 0/155
 - Arquitetura da Informação Quântica 0/2
 - Arquitetura de Cadeias de Suprimentos 0/1
 - ▣ Arquivologia 0/0
 - Repositórios Arquivísticos e seus Conteúdos 0/0
 - Técnicas e Organização dos Arquivos 0/0
 - ▣ Biblioteconomia 0/157
 - Ontologia 0/56
 - Processos de Disseminação da Informação 0/36
 - Técnicas de Recuperação da Informação 0/55
 - Técnicas de Tratamento Documental 0/12
 - Teorias da Classificação / Taxonomia 0/14
 - Fundamentos de Ciência da Informação 0/64
 - Gestão do Conhecimento 0/29
 - Processos de Comunicação 0/24
 - Profissional da Informação 0/2
 - Teoria da Informação 0/48
 - ▣ Comunicação 0/10
 - Teorias da Comunicação 0/9

Figura 3 – Classes 1 e 3

- ▣ Classe 1 0/230
 - ▣ Ciências Cognitivas 0/40
 - Psico-Linguística 0/3
 - Teorias da Ciência Cognitiva 0/34
 - ▣ Filosofia 0/162
 - Cosmologia 0/0
 - Empirismo 0/6
 - Epistemologia 0/26
 - Estruturalismo 0/3
 - Fenomenologia 0/60
 - Filosofia Analítica 0/6
 - Hermenêutica 0/20
 - Metodologia Científica 0/24
 - ▣ Lógica 0/22
 - ▣ Lógica Clássica 0/12
 - Lógica de Predicados 0/2
 - Lógica Proposicional 0/5
 - Lógica Descritiva 0/0
 - Lógica Quântica 0/0
 - ▣ Lógicas Modais 0/6
 - Lógicas da Sociedade 0/4
 - Lógicas Deônticas 0/1
 - Lógicas Doxásticas 0/1
 - Lógicas Epistêmicas 0/0
 - Lógicas Modais Normais 0/1
 - Lógicas Multimodais 0/0
 - Lógicas Nebulosas 0/0
 - Lógicas Paraconsistentes 0/3
 - Lógicas Polivalentes 0/0
 - ▣ Psicologia 0/18
 - Desenvolvimento e Capacidade Mental / Psicologia Comparada 0/12
 - Estudos e Processos Mentais Especiais 0/4
 - Métodos Psicológicos em Geral 0/1
 - Sensação e Percepção Sensorial 0/1
 - Sistemas e Escolas Psicológicas 0/0
- ▣ Classe 3 0/129
 - ▣ Administração 0/107
 - ▣ Educação 0/23
 - Administração Educacional 0/11
 - ▣ Fundamentos da Educação 0/14
 - Filosofia da Educação 0/0
 - Teoria da Educação 0/10

Figura 4 – Classe 5

- ▣ Classe 5 0/35
 - ▣ Biologia 0/13
 - Bases Materiais da Vida 0/1
 - Biologia Geral e Teórica 0/2
 - Ecologia Geral, Biotecnologia, Hidrobiologia, Biogeografia 0/2
 - Genética Geral, Citogenética Geral, Immunogenética, Evolução 0/5
 - ▣ Física 0/9
 - Aplicação/Utilização dos Fenômenos 0/0
 - Calor / Termodinâmica 0/2
 - ▣ Efeitos do Fenômenos 0/1
 - Efeitos Biológicos 0/0
 - Efeitos Físicos 0/0
 - Efeitos Químicos 0/0
 - Eleticidade / Magnetismo / Eletromagnetismo 0/0
 - Fórmulas Empíricas 0/0
 - Leis Gerais dos Fenômenos 0/0
 - Mecânica 0/1
 - Natureza Física da Matéria 0/0
 - Observação e Registros dos Fenômenos 0/0
 - Princípios Básicos da Física 0/0
 - Produção e Causa dos Fenômenos 0/0
 - Teoria e Natureza dos Fenômenos 0/3
 - ▣ Matemática 0/14
 - Álgebra 0/2
 - Análise Combinatória / Teoria dos Gráficos 0/1
 - Análise Matemática 0/2
 - Geometria 0/4
 - ▣ Reflexões Fundamentais e Gerais sobre Matemática 0/1
 - Fundamentos da Matemática 0/0
 - Teoria de Números 0/0
 - Topologia 0/0

Figura 5 – Classes 6,7 e 8

- ▣ Classe 6 0/157
 - ▣ Ciência da Computação 0/156
 - Complexidade 0/1
 - Computabilidade 0/76
 - ▣ Inteligência Artificial 0/37
 - Processamento de Linguagem Natural 0/1
 - Sistemas Baseados em Conhecimento 0/16
 - Teoria e Fundamentos da Inteligência Artificial 0/12
 - Processamento de Sinais 0/2
 - ▣ Programação em Lógica 0/5
 - Programação em Lógica Clássica 0/2
 - Programação em Lógica Não Clássica 0/0
 - Programação Procedimental 0/5
 - SGBD 0/22
 - Teoria da Computação 0/19
- ▣ Classe 7 0/13
 - ▣ Design 0/13
 - Design da Informação 0/0
 - Design de Interação 0/1
 - Design de Interface 0/4
 - Design de Produto 0/2
 - Design Gráfico 0/0
 - Fundamentos do Design 0/2
- ▣ Classe 8 0/56
 - ▣ Linguística 0/44
 - Análise do Discurso 0/23
 - Escolas e Tendências em Linguística 0/4
 - Gramática 0/3
 - Linguística Geral 0/5
 - Semântica 0/10
 - Sintaxe 0/0
 - ▣ Semiótica 0/13
 - Semiótica da Informação 0/6
 - Teoria e Fundamentos da Semiótica 0/7